Applications of AI – Image Processing
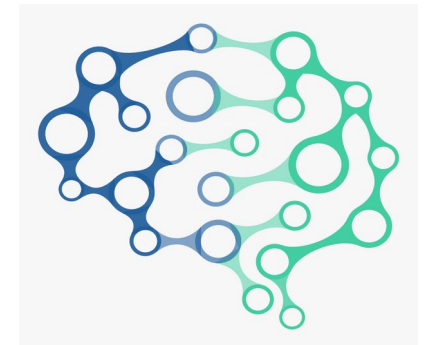
Tushar B. Kute,
http://tusharkute.com

Image Processing

- Generally speaking, image processing is manipulating an image in order to enhance it or extract information from it. There are two methods of image processing:

  – Analog image processing is used for processing physical photographs, printouts, and other hard copies of images

  – Digital image processing is used for manipulating digital images with the help of computer algorithms

Image Processing

- Here are some of the main purposes of image processing:
  - Visualization — Represent processed data in an understandable way, giving visual form to objects that aren't visible, for instance
  - Image sharpening and restoration — Improve the quality of processed images
  - Image retrieval — Help with image search
  - Object measurement — Measure objects in an image
  - Pattern recognition — Distinguish and classify objects in an image, identify their positions, and understand the scene
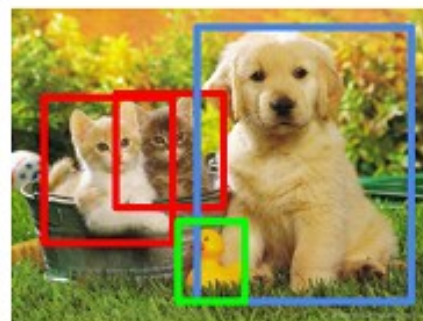
Image Processing

Image Processing

KEY PHASES OF DIGITAL IMAGE PROCESSING
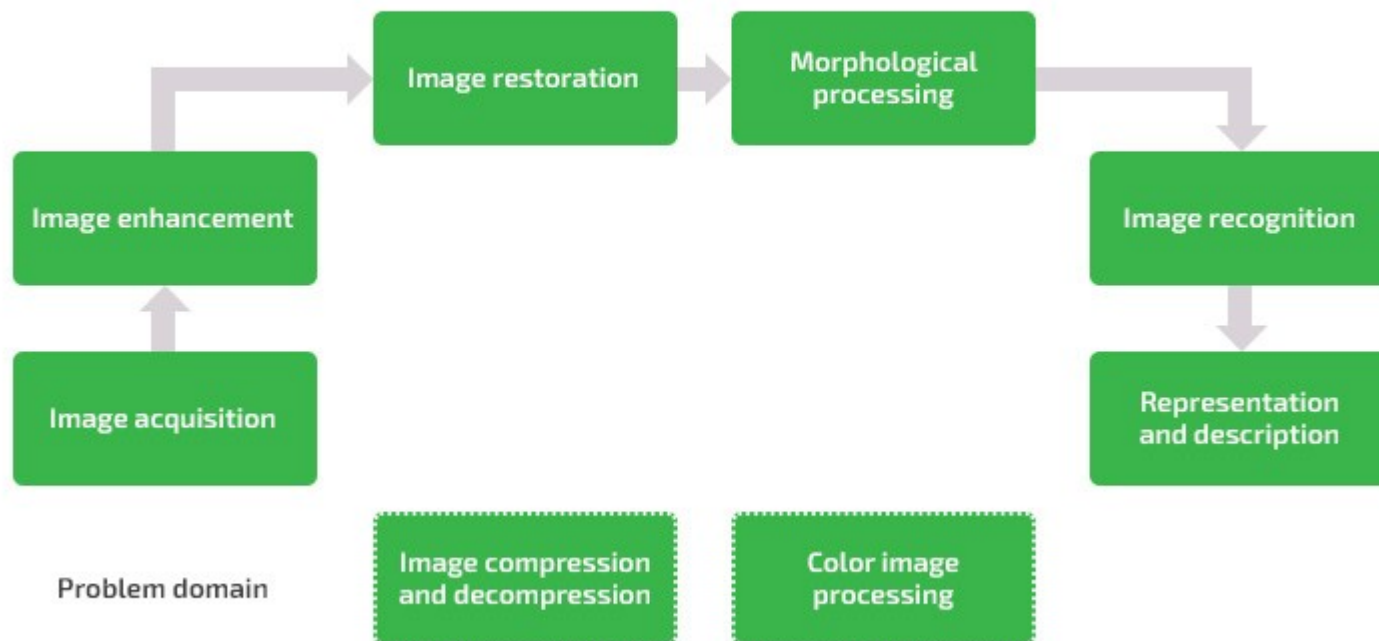
Image Processing

- Image acquisition is the process of capturing an image with a sensor (such as a camera) and converting it into a manageable entity (for example, a digital image file).

- One popular image acquisition method is scraping.

- Image enhancement improves the quality of an image in order to extract hidden information from it for further processing.

- Image restoration also improves the quality of an image, mostly by removing possible corruptions in order to get a cleaner version.

- This process is based mostly on probabilistic and mathematical models and can be used to get rid of blur, noise, missing pixels, camera misfocus, watermarks, and other corruptions that may negatively affect the training of a neural network.
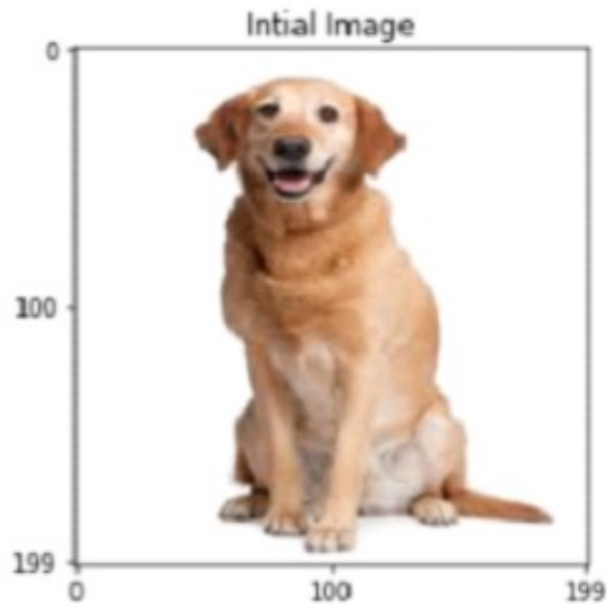
Image Processing

- Color image processing includes the processing of colored images and different color spaces. Depending on the image type, we can talk about pseudocolor processing (when colors are assigned grayscale values) or RGB processing (for images acquired with a full-color sensor).

- Image compression and decompression allow for changing the size and resolution of an image.

  – Compression is responsible for reducing the size and resolution, while decompression is used for restoring an image to its original size and resolution.

- These techniques are often used during the image augmentation process.

- When you lack data, you can extend your dataset with slightly augmented images. In this way, you can improve the way your neural network model generalizes data and make sure it provides high-quality results.

Image Processing : Augmentation

Augmented Images
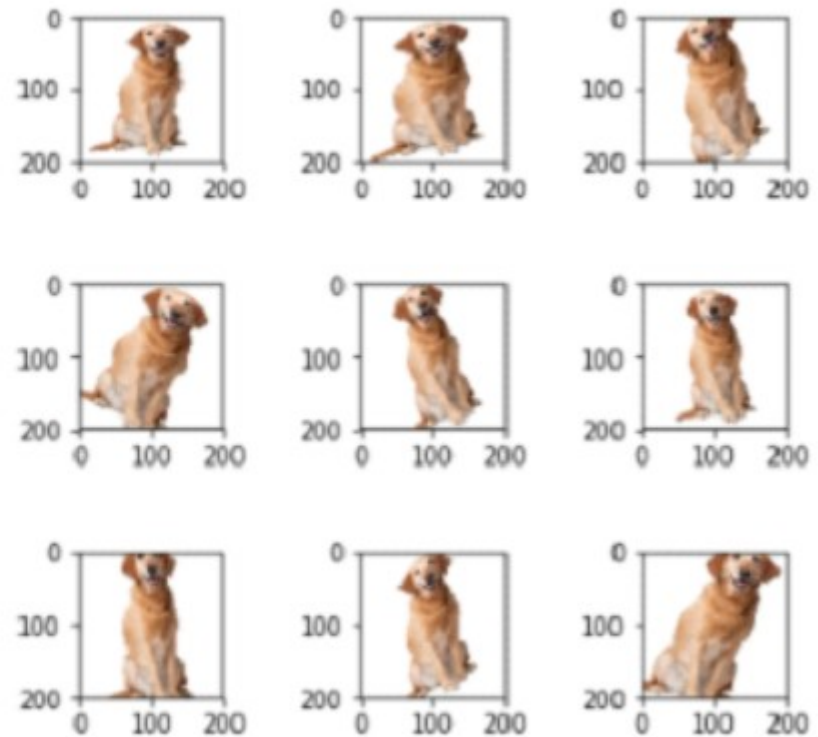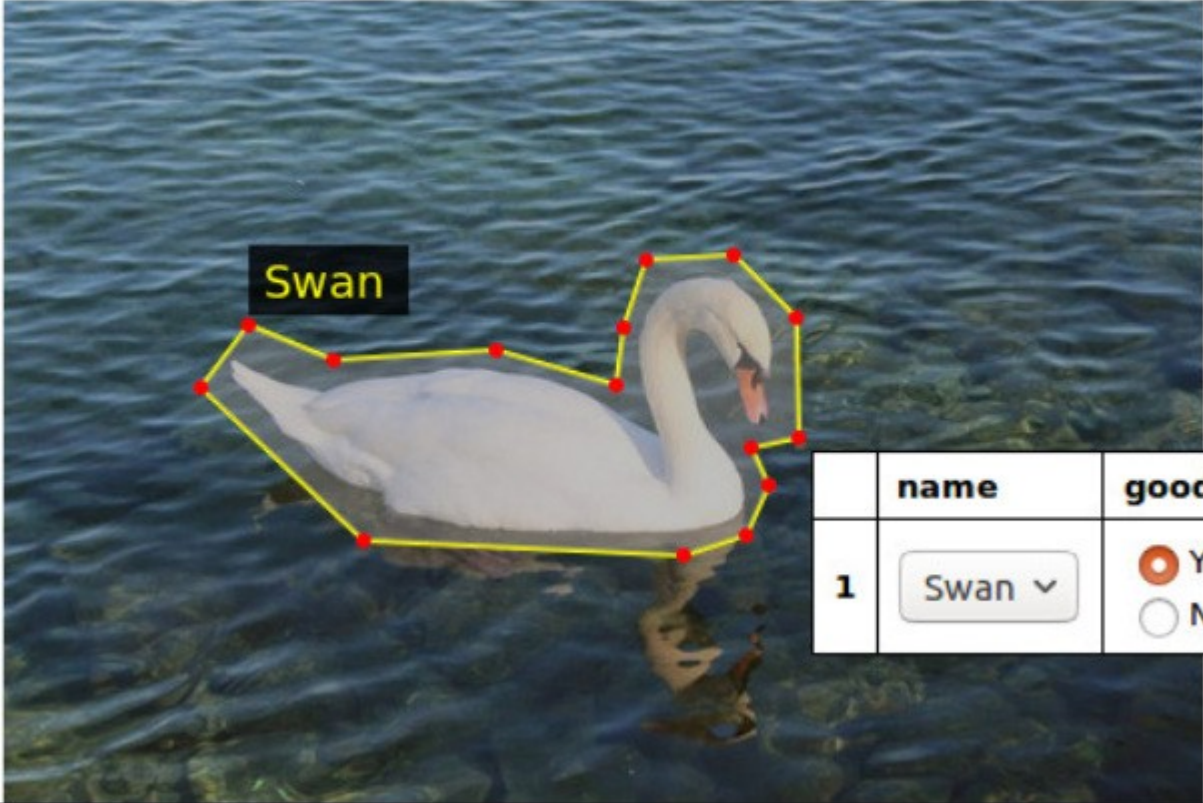
Image Processing

- Morphological processing describes the shapes and structures of the objects in an image.

- Morphological processing techniques can be used when creating datasets for training AI models.

- In particular, morphological analysis and processing can be applied at the annotation stage, when you describe what you want your AI model to detect or recognize.

Image Processing

- Image recognition is the process of identifying specific features of particular objects in an image.

- AI-based image recognition often uses such techniques as object detection, object recognition, and segmentation.

convolution + nonlinearity | max pooling | vec

convolution + pooling layers | fully connected layers | Nx binary classification

bird → P_bird
sunset → P_sunset
dog → P_dog
cat → P_cat
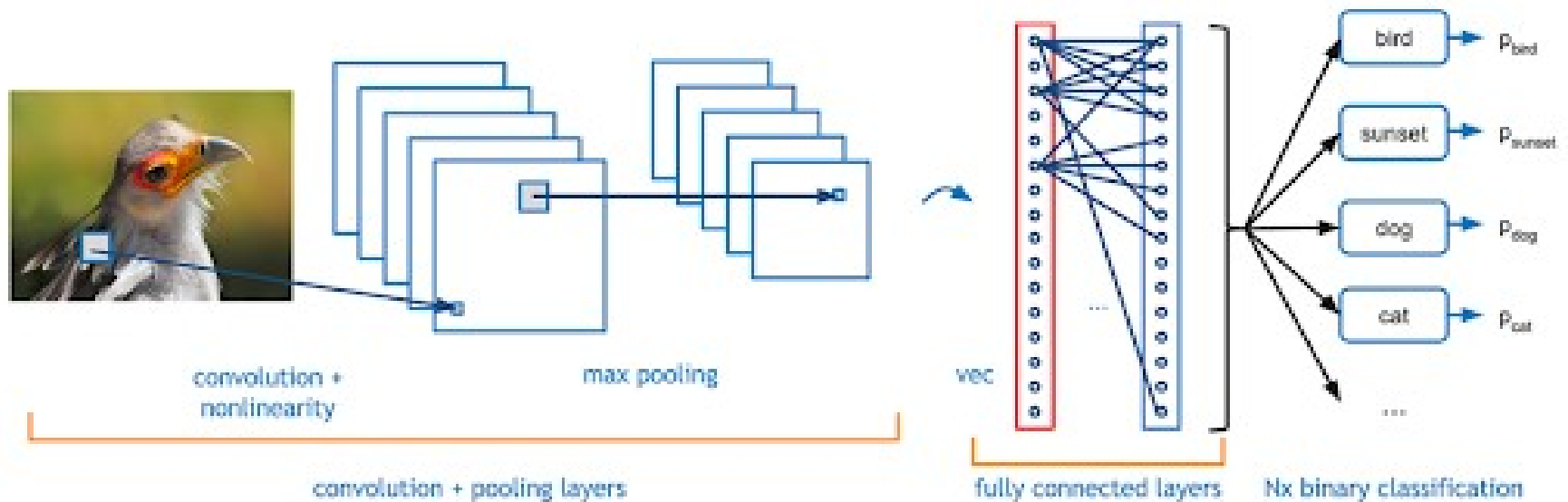
Image Processing

- Representation and description is the process of visualizing and describing processed data. AI systems are designed to work as efficiently as possible.

- The raw output of an AI system looks like an array of numbers and values that represent the information the AI model was trained to produce.

- Yet for the sake of system performance, a deep neural network usually doesn't include any output data representations.

- Using special visualization tools, you can turn these arrays of numbers into readable images suitable for further analysis.

Image Processing: Using AI/ML

- The use of AI and ML boosts both the speed of data processing and the quality of the final result.

- For instance, with the help of AI platforms, we can successfully accomplish such complex tasks as object detection, face recognition, and text recognition.

- But of course, in order to get high-quality results, we need to pick the right methods and tools for image processing.

tusharkute
.com

Image Processing

- Most images taken with regular sensors require preprocessing, as they can be misfocused or contain too much noise. Filtering and edge detection are two of the most common methods for processing digital images.

- Filtering is used for enhancing and modifying the input image. With the help of different filters, you can emphasize or remove certain features in an image, reduce image noise, and so on. Popular filtering techniques include linear filtering, median filtering, and Wiener filtering.

Image Processing

- Edge detection uses filters for image segmentation and data extraction. By detecting discontinuities in brightness, this method helps to find meaningful edges of objects in processed images.

- Canny edge detection, Sobel edge detection, and Roberts edge detection are among the most popular edge detection techniques.
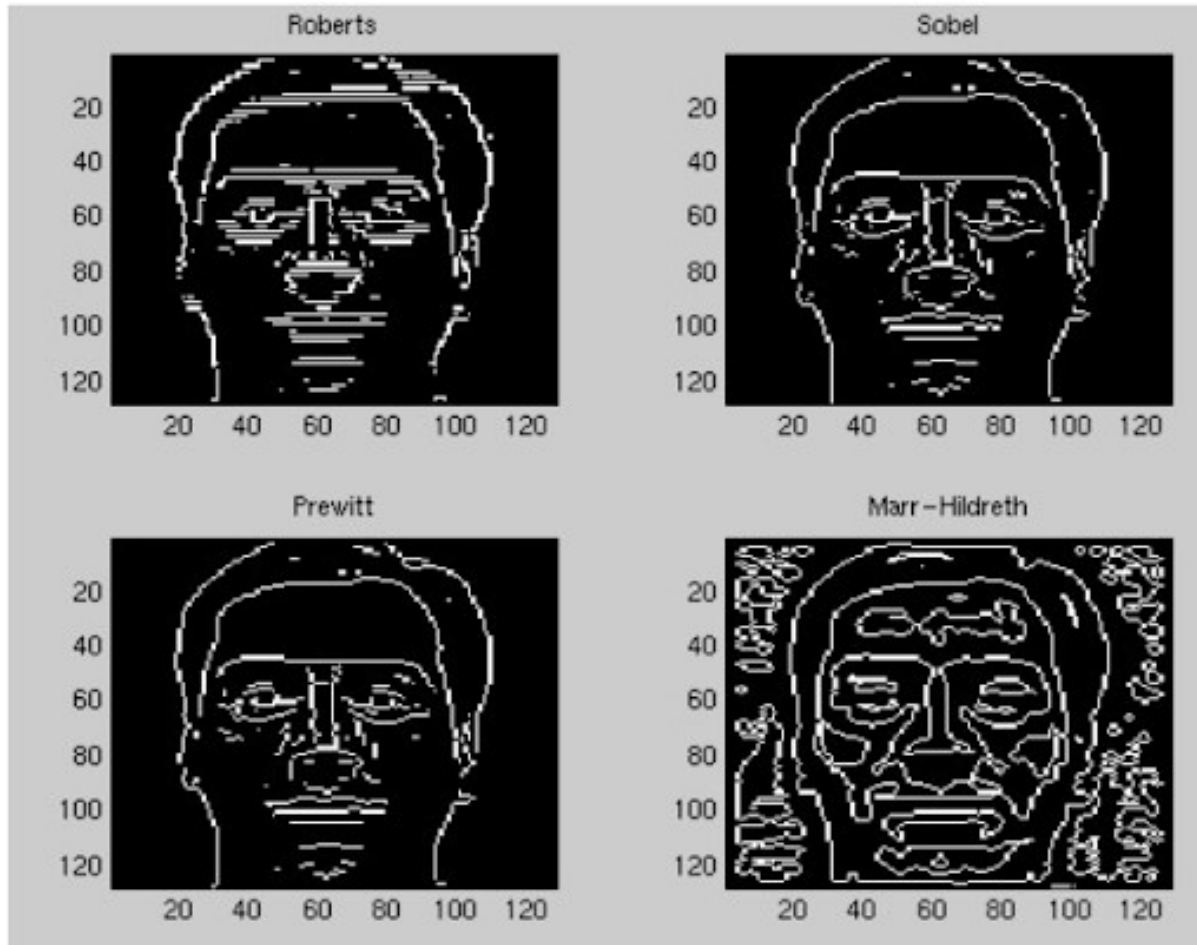
Image Processing

Image Processing

- There are also other popular techniques for handling image processing tasks. The wavelets technique is widely used for image compression, although it can also be used for denoising.

- Some of these filters can also be used as augmentation tools. For example, in one of our recent projects, we developed an AI algorithm that uses edge detection to discover the physical sizes of objects in digital image data.

- To make it easier to use these techniques as well as to implement AI-based image processing functionalities in your product, you can use specific libraries and frameworks. In the next section, we take a look at some of the most popular open-source libraries for accomplishing different image processing tasks with the help of AI algorithms.

- Computer vision libraries contain common image processing functions and algorithms.
- There are several open-source libraries you can use when developing image processing and computer vision features:
  - OpenCV
  - Visualization Library
  - VGG Image Annotator

OpenCV

- The Open Source Computer Vision Library (OpenCV) is a popular computer vision library that provides hundreds of computer and machine learning algorithms and thousands of functions composing and supporting those algorithms.

- The library comes with C++, Java, and Python interfaces and supports all popular desktop and mobile operating systems.

- OpenCV includes various modules, such as an image processing module, object detection module, and machine learning module.

- Using this library, you can acquire, compress, enhance, restore, and extract data from images.

Visualization Library

- Visualization Library is C++ middleware for 2D and 3D applications based on the Open Graphics Library (OpenGL).

- This toolkit allows you to build portable and high-performance applications for Windows, Linux, and Mac OS X systems.

- As many of the Visualization Library classes have intuitive one-to-one mapping with functions and features of the OpenGL library, this middleware is easy and comfortable to work with.

VGG Image Annotator

- VGG Image Annotator (VIA) is a web application for object annotation.

- It can be installed directly in a web browser and used for annotating detected objects in images, audio, and video records.

- VIA is easy to work with, doesn't require additional setup or installation, and can be used with any modern browser.

Machine Learning Frameworks

- If you want to move beyond using simple AI algorithms, you can build custom deep learning models for image processing. To make development a bit faster and easier, you can use special platforms and frameworks. Below, we take a look at some of the most popular ones:
  - TensorFlow
  - PyTorch
  - MATLAB Image Processing Toolbox
  - Microsoft Computer Vision
  - Google Cloud Vision
  - Google Colaboratory (Colab)

Tensorflow

- Google's TensorFlow is a popular open-source framework with support for machine learning and deep learning.

- Using TensorFlow, you can create and train custom deep learning models.

- The framework also includes a set of libraries, including ones that can be used in image processing projects and computer vision applications.

PyTorch

- PyTorch is an open-source deep learning framework initially created by the Facebook AI Research lab (FAIR).

- This Torch-based framework has Python, C++, and Java interfaces.

- Among other things, you can use PyTorch for building computer vision and natural language processing applications.

Matlab

- MATLAB is an abbreviation for matrix laboratory. It's the name of both a popular platform for solving scientific and mathematical problems and a programming language.

- This platform provides an Image Processing Toolbox (IPT) that includes multiple algorithms and workflow applications for AI-based picture analysis, processing, and visualizing as well as for developing algorithms.

- MATLAB IPT allows you to automate common image processing workflows. This toolbox can be used for noise reduction, image enhancement, image segmentation, 3D image processing, and other tasks.

Microsoft Computer Vision

- Computer Vision is a cloud-based service provided by Microsoft that gives you access to advanced algorithms for image processing and data extraction. It allows you to:
  - analyze visual features and characteristics of an image
  - moderate image content
  - extract text from images

Google Cloud Vision

- Cloud Vision is part of the Google Cloud platform and offers a set of image processing features.

- It provides an API for integrating such features as image labeling and classification, object localization, and object recognition.

- Cloud Vision allows you to use pre-trained machine learning models and create and train custom models for creating image processing projects using machine learning.

Google Colab

- Google Colaboratory, otherwise known as Colab, is a free cloud service that can be used not only for improving your coding skills but also for developing deep learning applications from scratch.

- Colab makes it easier to use popular libraries such as OpenCV, Keras, and TensorFlow when developing an AI-based application.

- The service is based on Jupyter Notebooks, allowing AI developers to share their knowledge and expertise in a comfortable way. Plus, in contrast to similar services, Colab provides free GPU resources.

Neural Networks

- Most effective machine learning models for image processing use neural networks and deep learning. Deep learning uses neural networks for solving complex tasks similarly to the way the human brain solves them.

- Different types of neural networks can be deployed for solving different image processing tasks, from simple binary classification (whether an image does or doesn't match a specific criteria) to instance segmentation.

- Choosing the right type and architecture of a neural network plays an essential part in creating an efficient AI-based image processing solution.

tusharkute
.com

CNN

- Convolutional Neural Networks (ConvNets or CNNs) are a class of deep learning networks that were created specifically for image processing with AI.

- However, CNNs have been successfully applied on various types of data, not only images. In these networks, neurons are organized and connected similarly to how neurons are organized and connected in the human brain.

- In contrast to other neural networks, CNNs require fewer preprocessing operations. Plus, instead of using hand-engineered filters (despite being able to benefit from them), CNNs can learn the necessary filters and characteristics during training.

CNN

- All CNN layers are organized in three dimensions (weight, height, and depth) and have two components:
    - Feature extraction
    - Classification
- In the first component, the CNN runs multiple convolutions and pooling operations in order to detect features it will then use for image classification.

CNN

- CNNs are widely used for implementing AI in image processing and solving such problems as signal processing, image classification, and image recognition.

- There are numerous types of CNN architectures such as AlexNet, ZFNet, Faster R-CNN, and GoogLeNet/Inception.

- The choice of CNN architecture depends on the task at hand. For instance, GoogLeNet shows a higher accuracy for leaf recognition than AlexNet or a basic CNN. At the same time, due to the higher number of layers, GoogLeNet takes longer to run.

Mask R-CNN

- Mask R-CNN is a Faster R-CNN-based deep neural network that can be used for separating objects in a processed image or video. This neural network works in two stages:

  – Segmentation – The neural network processes an image, detects areas that may contain objects, and generates proposals.

  – Generation of bounding boxes and masks – The network calculates a binary mask for each class and generates the final results based on these calculations.

Mask R-CNN

Mask R-CNN

- Mask R-CNN remains one of the best solutions for instance segmentation.

- We have applied this neural network architecture and our image processing skills to solve many complex tasks, including the processing of medical image data and medical microscopic data.

- We've also developed a plugin for improving the performance of this neural network model up to ten times thanks to the use of NVIDIA TensorRT technology.

Fully CNN

- The concept of a fully convolutional network (FCN) was first offered by a team of researchers from the University of Berkeley.

- The main difference between a CNN and FCN is that the latter has a convolutional layer instead of a regular fully connected layer.

- As a result, FCNs are able to manage different input sizes. Also, FCNs use downsampling (striped convolution) and upsampling (transposed convolution) to make convolution operations less computationally expensive.

U-Net

- U-Net is a convolutional neural network that allows for fast and precise image segmentation. In contrast to other neural networks on our list, U-Net was designed specifically for biomedical image segmentation.

- Therefore, it comes as no surprise that U-Net is believed to be superior to Mask R-CNN especially in such complex tasks as medical image processing.

- U-Net has a U-shaped architecture and has more feature channels in its upsampling part. As a result, the network propagates context information to higher-resolution layers, thus creating a more or less symmetric expansive path to its contracting part.

U-Net

conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1
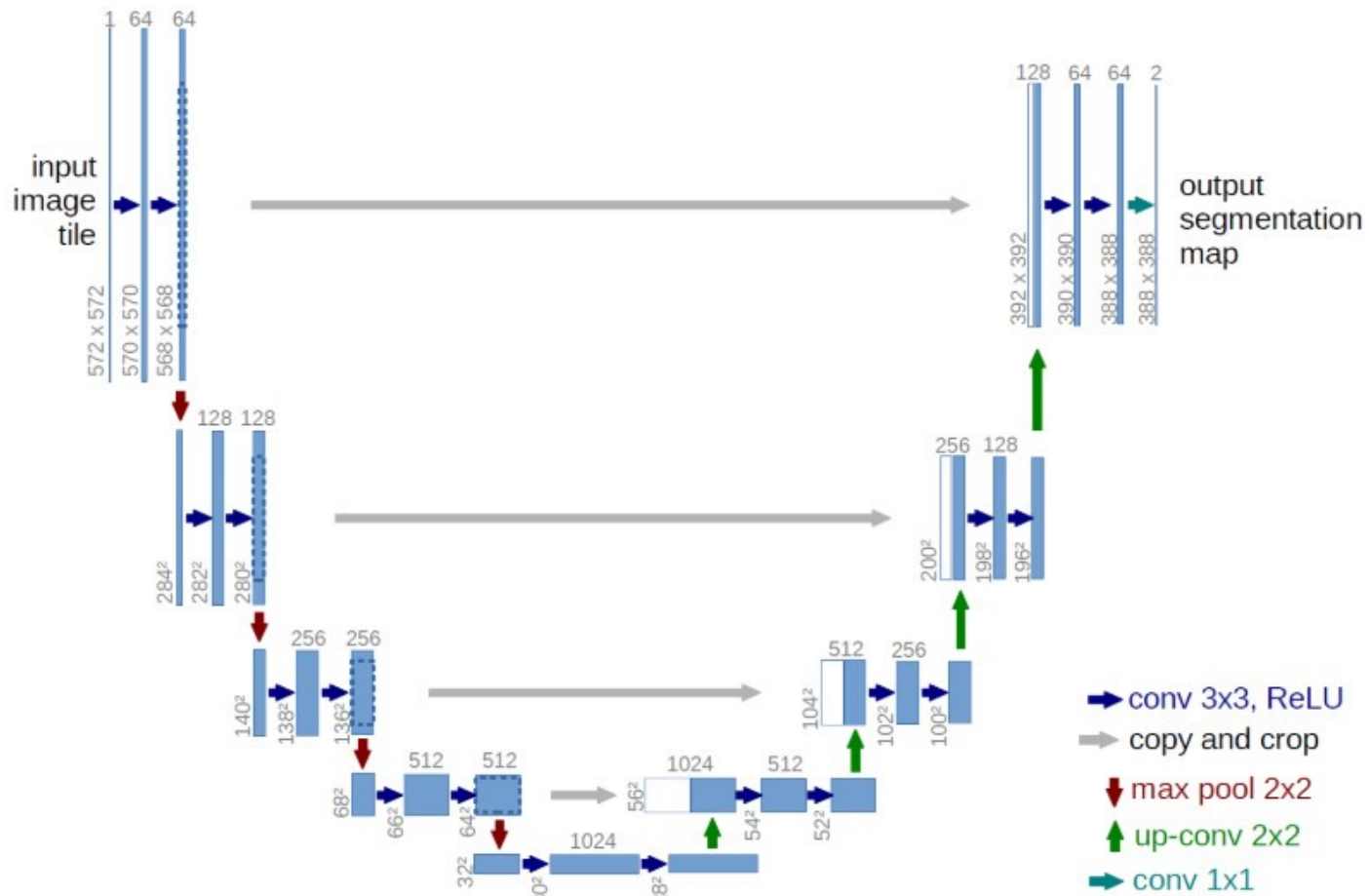
GAN

- Generative adversarial networks (GANs) are supposed to deal with one of the biggest challenges neural networks face these days: adversarial images.

- Adversarial images are known for causing massive failures in neural networks. For instance, a neural network can be fooled if you add a layer of visual noise called perturbation to the original image.

- And even though the difference is nearly unnoticeable to the human brain, computer algorithms struggle to properly classify adversarial images

GAN

"panda"
57.7% confidence

$+ \epsilon$

$=$

"gibbon"
99.3% confidence

Conclusion

- With the help of deep learning algorithms and neural networks, machines can be taught to see and interpret images in the way required for a particular task.

- Progress in the implementation of AI algorithms for image processing is impressive and opens a wide range of opportunities in fields from medicine and agriculture to retail and law enforcement.

Thank you

@mitu_skillologies

/mITuSkillologies

@mitu_group

/company/mitu-skillologies

MITUSkillologies

Web Resources
http://mitu.co.in
http://tusharkute.com

contact@mitu.co.in

tushar@tusharkute.com